DOCUMENT RESUME

ED 398 236 TM 024 736

AUTHOR Kehoe, Jerard
TITLE Writing Multiple-Choice Test Items. ERIC/AE Digest.
INSTITUTION ERIC Clearinghouse on Assessment and Evaluation, Washington, DC.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO EDO-TM-95-3
PUB DATE Oct 95
CONTRACT RR93002002
NOTE 4p.; Adapted from "Testing Memo 4: Constructing Multiple-Choice Tests--Part I," Office of Measurement and Research Services, Virginia Polytechnic Institute and State University, Blacksburg, VA 24060.
AVAILABLE FROM ERIC Clearinghouse on Assessment and Evaluation, The Catholic University of America, Department of Education, O'Boyle Hall, Washington, DC 20064 (free).
PUB TYPE Information Analyses - ERIC Clearinghouse Products (071)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Culture Fair Tests; *Distractors (Tests); Educational Assessment; Item Bias; Measurement Techniques; *Multiple Choice Tests; Scoring; *Statistical Analysis; Stereotypes; *Test Construction; Test Items; Test Theory
IDENTIFIERS ERIC Digests; *Item Writing

ABSTRACT

This digest describes some basics of the construction of multiple-choice tests. As a rule, the test maker should strive for test item stems (introductory questions or incomplete statements at the beginning of each item that are followed by the options) that are clear and parsimonious, answers that are unequivocal and chosen by the students who do best on the test, and distractors that are plausible competitors for the item as evidenced by the frequency with which they are chosen. It is very important that items be developed over time in light of the evidence that can be obtained from the statistical output of a measurement services office and from expert editorial review. Planning the general test outline is followed by writing item stems that identify the point to be tested and state the question in clear terms that do not lead to stereotyped answers. Negatives in the stem and irrelevant clues should be avoided. Writing the answer options calls for the same attention to clarity. Some specifics for option design and placement are provided. (Contains 4 references.) (SLD)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED 398 236

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

# WRITING MULTIPLE-CHOICE TEST ITEMS

ERIC/AE Digest Series EDO-TM-95-3
The Catholic University of America

Jerard Kehoe

BEST COPY AVAILABLE

TM024736

CUA

# ERIC/AE DIGEST

DEPARTMENT OF EDUCATION THE CATHOLIC UNIVERSITY OF AMERICA

EDO-TM-95-3 October 1995

## Writing Multiple-Choice Test Items

*Jerard Kehoe*
*Virginia Polytechnic Institute and State University*

A notable concern of many teachers is that they frequently have the task of constructing tests but have relatively little training or information to rely on in this task. The objective of this Digest is to set out some conventional wisdom for the construction of multiple-choice tests, which are one of the most common forms of teacher-constructed tests. The comments which follow are applicable mainly to multiple-choice tests covering fairly broad topic areas.

Before proceeding, it will be useful to establish our terms for discussing multiple-choice items. The *stem* is the introductory question or incomplete statement at the beginning of each item and this is followed by the options. The *options* consist of the answer -- the correct option -- and *distractors*--the incorrect but (we hope) tempting options.

### General Objectives

As a rule, one is concerned with writing stems that are clear and parsimonious, answers that are unequivocal and chosen by the students who do best on the test, and distractors that are plausible competitors of the answer as evidenced by the frequency with which they are chosen. Lastly, and probably most important, we should adopt the attitude that items need to be developed over time in the light of evidence that can be obtained from the statistical output typically provided by a measurement services office (where tests are machine-scored) and from "expert" editorial review.

### Planning

The primary objective in planning a test is to outline the actual course content that the test will cover. A convenient way of accomplishing this is to take 10 minutes following each class to list on an index card the important concepts covered in class and in assigned reading for that day. These cards can then be used later as a source of test items. An even more conscientious approach, of course, would be to construct the test items themselves after each class. The advantage of either of these approaches is that the resulting test is likely to be a better representation of course activity than if the test were constructed before the course or after the course, when we usually have only a fond memory or optimistic syllabus to draw from. When we are satisfied that we have an accurate description of the content areas, then all that remains is to construct items that represent specific content areas. In developing good multiple-choice items, three tasks need to be considered: writing stems, writing options, and ongoing item development. The first two are discussed in this Digest.

### Writing Stems

We will first describe some basic rules for the construction of multiple-choice stems, because they are typically, though not necessarily, written before the options.

1. Before writing the stem, identify the one point to be tested by that item. In general, the stem should not pose more than one problem, although the solution to that problem may require more than one step.
2. Construct the stem to be either an incomplete statement or a direct question, avoiding stereotyped phraseology, as rote responses are usually based on verbal stereotypes. For example, the following stems (with answers in parentheses) illustrate undesirable phraseology:

   *What is the biological theory of recapitulation? (Ontogeny repeats phylogeny)*

   *Who was the chief spokesman for the "American System?" (Henry Clay)*

   Correctly answering these questions likely depends less on understanding than on recognizing familiar phraseology.
3. Avoid including nonfunctional words that do not contribute to the basis for choosing among the options. Often an introductory statement is included to enhance the appropriateness or significance of an item but does not affect the meaning of the problem in the item. Generally, such superfluous phrases should be excluded. For example, consider:

   *The American flag has three colors. One of them is (1) red (2) green (3) black*

   versus

   *One of the colors of the American flag is (1) red (2) green (3) black*



TM024736

ERIC Clearinghouse on Assessment and Evaluation

The Catholic University of America, Department of Education, O'Boyle Hall, Washington, DC 20064 202 319-5120

In particular, irrelevant material should not be used to make the answer less obvious. This tends to place too much importance on reading comprehension as a determiner of the correct option.

4. Include as much information in the stem and as little in the options as possible. For example, if the point of an item is to associate a term with its definition, the preferred format would be to present the definition in the stem and several terms as options rather than to present the term in the stem and several definitions as options.

5. Restrict the use of negatives in the stem. Negatives in the stem usually require that the answer be a false statement. Because students are likely in the habit of searching for true statements, this may introduce an unwanted bias.

6. Avoid irrelevant clues to the correct option. Grammatical construction, for example, may lead students to reject options which are grammatically incorrect as the stem is stated. Perhaps more common and subtle, though, is the problem of common elements in the stem and in the answer. Consider the following item:

   *What led to the formation of the States' Rights Party?*

   a. *The level of federal taxation*
   b. *The demand of states for the right to make their own laws*
   c. *The industrialization of the South*
   d. *The corruption of federal legislators on the issue of state taxation*

   One does not need to know U.S. history in order to be attracted to the answer, b.

Other rules that we might list are generally commonsensical, including recommendations for independent and important items and prohibitions against complex, imprecise wording.

**Writing Options**

Following the construction of the item stem, the likely more difficult task of generating options presents itself. The rules we list below are not likely to simplify this task as much as they are intended to guide our creative efforts.

1. Be satisfied with three or four well constructed options. Generally, the minimal improvement to the item due to that hard-to-come-by fifth option is not worth the effort to construct it. Indeed, all else the same, a test of 10 items each with four options is likely a better test than a test with nine items of five options each.

2. Construct distractors that are comparable in length, complexity and grammatical form to the answer, avoiding the use of such words as "always," "never," and "all." Adherence to this rule avoids some of the more common sources of biased cueing. For example, we sometimes find ourselves increasing the length and specificity of the answer (relative to distractors) in order to insure its truthfulness. This, however, becomes an easy-to-spot clue for the testwise student. Related to this issue is the question of whether or not test writers should take advantage of these types of cues to construct more tempting distractors. Surely not! The number of students choosing a distractor should depend only on deficits in the content area which the item targets and should not depend on cue biases or reading comprehension differences in "favor" of the distractor.

3. Options which read "none of the above," "both a. and e. above," "all of the above," _etc_., should be avoided when the students have been instructed to choose "the best answer," which implies that the options vary in degree of correctness. On the other hand, "none of the above" is acceptable if the question is factual and is probably desirable if computation yields the answer. "All of the above" is never desirable, as one recognized distractor eliminates it and two recognized answers identify it.

4. After the options are written, vary the location of the answer on as random a basis as possible. A convenient method is to flip two (or three) coins at a time where each possible Head-Tail combination is associated with a particular location for the answer. Furthermore, if the test writer is conscientious enough to randomize the answer locations, students should be informed that the locations are randomized. (Testwise students know that for some instructors the first option is rarely the answer.)

5. If possible, have a colleague with expertise in the content area of the exam review the items for possible ambiguities, redundancies or other structural difficulties. Having completed the items we are typically so relieved that we may be tempted to regard the task as completed and each item in its final and permanent form. Yet, another source of item and test improvement is available to us, namely, statistical analyses of student responses.

This Digest was adapted with permission from *Testing Memo 4: Constructing Multiple-Choice Tests -- Part I*, Office of Measurement and Research Services, Virginia Polytechnic Institute and State University, Blacksburg, VA 24060

**Further Reading**

Airasian, P. (1994) *Classroom Assessment*, Second Edition, NY: McGraw-Hill.

Cangelosi, J. (1990) *Designing Tests for Evaluating Student Achievement*. NY: Addison Wellesley.

Grunlund, N (1993) *How to make achievement tests and assessments*, 5th edition, NY: Allen and Bacon.

Haladyna, T.M. & Downing, S.M. (1989) Validity of a Taxonomy of Multiple-Choice Item-Writing Rules. *Applied Measurement in Education*, 2 (1), 51-78.

BEST COPY AVAILABLE



This publication was prepared with funding from the Office of Educational Research and Improvement, U.S. Department of Education, under contract RR93002002. The opinions expressed in this report do not necessarily reflect the positions or policies of OERI or the U.S. Department of Education. **Permission is granted to copy and distribute this ERIC/AE Digest.**